

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS: BEAUDOIN, Adrien R.
SEVIGNY, Jean

(ii) TITLE OF THE INVENTION: ATP-DIPHOSPHOHYDROLASES, PROCESS
OF PURIFICATION THEREOF AND PROCESS OF PRODUCING THEREOF BY
RECOMBINANT TECHNOLOGY

(iii) NUMBER OF SEQUENCES: 8

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: GOUDREAU GAGE DUBUC & MARTINEAU WALKER
(B) STREET: 800 PLACE VICTORIA, SUITE 3400
(C) CITY: MONTREAL
(D) STATE: QUEBEC
(E) COUNTRY: CANADA
(F) ZIP: H4Z 1E9

(v) COMPUTER READABLE :

(A) MEDIUM TYPE: DISKETTE 1.44
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: WINDOWS 3.1
(D) SOFTWARE: WORDPERFECT 6.1

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE: APRIL 10, 1996
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION: U.S.S.N 08/419,204
(B) FILING DATE: APRIL 10, 1995

(viii) ATTORNEY/AGENT INFORMATION:

P021P00-530278260

(A) NAME: Jean H. DUBUC, Gaétan PRINCE, Alain M.
LECLERC
(C) REFERENCE/DOCKET NUMBER: DH/10857.146

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (514) 397-4335
(B) TELEFAX: (514) 397-4382

(2) INFORMATION FOR SEQ ID NO: 1

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 510 AMINO ACIDS
(B) TYPE: AMINO ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1

Met Glu Asp Thr Lys Glu Ser Asn Val Lys Thr Phe Cys Ser Lys
5 10 15
Asn Ile Leu Ala Ile Leu Gly Phe Ser Ser Ile Ile Ala Val Ile
20 25 30
Ala Leu Leu Ala Val Gly Leu Thr Gln Asn Lys Ala Leu Pro Glu
35 40 45
Asn Val Lys Tyr Gly Ile Val Leu Asp Ala Gly Ser Ser His Thr
50 55 60
Ser Leu Tyr Ile Tyr Lys Trp Pro Ala Glu Lys Glu Asn Asp Thr
65 70 75
Gly Val Val His Gln Val Glu Glu Cys Arg Val Lys Gly Pro Gly
80 85 90
Ile Ser Lys Phe Val Gln Lys Val Asn Glu Ile Gly Ile Tyr Leu
95 100 105
Thr Asp Cys Met Glu Arg Ala Arg Glu Val Ile Pro Arg Ser Gln
110 115 120
His Gln Glu Thr Pro Val Tyr Leu Gly Ala Thr Ala Gly Met Arg
125 130 135
Leu Leu Arg Met Glu Ser Glu Glu Leu Ala Asp Arg Val Leu Asp
140 145 150
Val Val Glu Arg Ser Leu Ser Asn Tyr Pro Phe Asp Phe Gln Gly
155 160 165

Ala Arg Ile Ile Thr Gly Gln Glu Glu Gly Ala Tyr Gly Trp Ile
170 175 180
Thr Ile Asn Tyr Leu Leu Gly Lys Phe Ser Gln Lys Thr Arg Trp
185 190 195
Phe Ser Ile Val Pro Tyr Glu Thr Asn Asn Gln Glu Thr Phe Gly
200 205 210
Ala Leu Asp Leu Gly Gly Ala Ser Thr Gln Val Thr Phe Val Pro
215 220 225
Gln Asn Gln Thr Ile Glu Ser Pro Asp Asn Ala Leu Gln Phe Arg
230 235 240
Leu Tyr Gly Lys Asp Tyr Asn Val Tyr Thr His Ser Phe Leu Cys
245 250 255
Tyr Gly Lys Asp Gln Ala Leu Trp Gln Lys Leu Ala Lys Asp Ile
260 265 270
Gln Val Ala Ser Asn Glu Ile Leu Arg Asp Pro Cys Phe His Pro
275 280 285
Gly Tyr Lys Lys Val Val Asn Val Ser Asp Leu Tyr Lys Thr Pro
290 295 300
Cys Thr Lys Arg Phe Glu Met Thr Leu Pro Phe Gln Gln Phe Glu
305 310 315
Ile Gln Gly Ile Gly Asn Tyr Gln Gln Cys His Gln Ser Ile Leu
320 325 330
Glu Leu Phe Asn Thr Ser Tyr Cys Pro Tyr Ser Gln Cys Ala Phe
335 340 345
Asn Gly Ile Phe Leu Pro Pro Leu Gln Gly Asp Phe Gly Ala Phe
350 355 360
Ser Ala Phe Tyr Phe Val Met Lys Phe Leu Asn Leu Thr Ser Glu
365 370 375
Lys Val Ser Gln Glu Lys Val Thr Glu Met Met Lys Phe Cys
380 385 390
Ala Gln Pro Trp Glu Glu Ile Lys Thr Ser Tyr Ala Gly Val Lys
395 400 405
Glu Lys Tyr Leu Ser Glu Tyr Cys Phe Ser Gly Thr Tyr Ile Leu
410 415 420
Ser Leu Leu Leu Gln Gly Tyr His Phe Thr Ala Asp Ser Trp Glu
425 430 435
His Ile His Phe Ile Gly Lys Ile Gln Gly Ser Asp Ala Gly Trp
440 445 450
Thr Leu Gly Tyr Met Leu Asn Leu Thr Asn Met Ile Pro Ala Glu
455 460 465
Gln Pro Leu Ser Thr Pro Leu Ser His Ser Thr Tyr Val Phe Leu
470 475 480

Met Val Leu Phe Ser Leu Val Leu Phe Thr Val Ala Ile Ile Gly
485 490 495
Leu Leu Ile Phe His Lys Pro Ser Tyr Phe Trp Lys Asp Met Val
500 505 510

(2) INFORMATION FOR SEQ ID NO: 2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1818 BASE PAIRS
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2

ACACACCAA GCAGCGGCTG GGGGGGGAA AGACGAGGAA AGAGGAGGAA 50
AACAAAAGCT GCTACTTATG GAAGATACAA AGGAGTCTAA CGTGAAGACA 100
TTTGCTCCA AGAATATCCT AGCCATCCTT GGCTTCTCCT CTATCATAGC 150
TGTGATAGCT TTGCTTGCTG TGGGTTGAC CCAGAACAAA GCATTGCCAG 200
AAAACGTTAA GTATGGGATT GTGCTGGATG CGGGTTCTTC TCACACAAAGT 250
TTATACATCT ATAAGTGGCC AGCAGAAAAG GAGAATGACA CAGGCGTGGT 300
GCATCAAGTA GAAGAATGCA GGGTTAAAGG TCCTGGAATC TCAAAATTG 350
TTCAGAAAGT AAATGAAATA GGCATTACCG TGACTGATTG CATGGAAAGA 400
GCTAGGGAAAG TGATTCCAAG GTCCCAGCAC CAAGAGACAC CCGTTTACCT 450
GGGAGGCCACG GCAGGCATGC GGTTGCTCAG GATGGAAAGT GAAGAGTTGG 500
CAGACAGGGT TCTGGATGTG GTGGAGAGGA GCCTCAGCAA CTACCCCTTT 550
GACTTCCAGG GTGCCAGGAT CATTACTGGC CAAGAGGAAG GTGCCTATGG 600
CTGGATTACT ATCAACTATC TGCTGGCAA ATTCACTGAG AAAACAAGGT 650
GGTCAGCAT AGTCCCCATAT GAAACCAATA ATCAGGAAAC CTTTGGAGCT 700
TTGGACCTTG GGGGAGCCTC TACACAAGTC ACTTTGTAC CCCAAAACCA 750
GACTATCGAG TCCCCAGATA ATGCTCTGCA ATTTCGCCTC TATGGCAAGG 800
ACTACAATGT CTACACACAT AGCTTCTTGT GCTATGGAA GGATCAGGCA 850
CTCTGGCAGA AACTGGCAA GGACATTCAG GTTGCAGTA ATGAAATTCT 900
CAGGGACCCA TGCTTTCATC CTGGATATAA GAAGGTAGTG AACGTAAGTG 950
ACCTTTACAA GACCCCTGC ACCAAGAGAT TTGAGATGAC TCTTCCATTG 1000
CAGCAGTTG AAATCCAGGG TATTGAAAC TATCAACAAT GCCATCAAAG 1050
CATCCTGGAG CTCTTCAACA CCAGTTACTG CCCTTACTCC CAGTGTGCCT 1100
TCAATGGGAT TTTCTTGCCA CCACTCCAGG GGGATTTGG GGCATTTCA 1150
GCTTTTACT TTGTGATGAA GTTTTAAAC TTGACATCAG AGAAAGTCTC 1200
TCAGGAAAAG GTGACTGAGA TGATGAAAAA GTTCTGTGCT CAGCCTGGG 1250
AGGAGATAAA AACATCTTAC GCTGGAGTAA AGGAGAAGTA CCTGAGTGAA 1300

TACTGCTTT	CTGGTACCTA	CATTCTCTCC	CTCCTTCTGC	AAGGCTATCA	1350
TTTCACAGCT	GATTCTGGG	AGCACATCCA	TTTCATTGGC	AAGATCCAGG	1400
GCAGCGACGC	CGGCTGGACT	TTGGGCTACA	TGCTGAACCT	GACCAACATG	1450
ATCCCAGCTG	AGCAACCATT	GTCCACACCT	CTCTCCCCT	CCACCTATGT	1500
CTTCCTCATG	GTTCTATTCT	CCCTGGTCT	TTTCACAGTG	GCCATCATAG	1550
GCTTGCTTAT	CTTTCACAAG	CCTTCATATT	TCTGGAAAGA	TATGGTATAG	1600
CAAAAGCAGC	TGAAATATGC	TGGCTGGAGT	GAGGAAAAAA	TCGTCCAGGG	1650
AGCATTTC	TCCATCGCAG	TGTTCAAGGC	CATCCTTCCC	TGTCTGCCAG	1700
GGCCAGTCTT	GACGAGTGTG	AAGCTTCCTT	GGCTTTACT	GAAGCCTTTC	1750
TTTTGGAGGT	ATTCAATATC	CTTTGCCTCA	AGGACTTCGG	CAGATACTGT	1800
CTCTTCATG	AGTTTTTC				1818

(2) INFORMATION FOR SEQ ID NO: 3

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3

Glu Thr Pro Val Tyr Leu Gly Ala Thr Ala Gly
5 10

(2) INFORMATION FOR SEQ ID NO: 4

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4

Leu Leu Arg Met Glu
5

(2) INFORMATION FOR SEQ ID NO: 5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 AMINO ACIDS

(B) TYPE: AMINO ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5

Ala Asp Lys Ile Leu Ala Asn Xaa Val Ala Ser Ser Ile
Ala Asp Lys Ile Leu Ala Asn Xaa Val Ala Ser Ser Ile
5 10

(2) INFORMATION FOR SEQ ID NO: 6

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 AMINO ACIDS

(B) TYPE: AMINO ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6

Tyr Pro Phe Asp Phe Gln Gly Ala Arg Ile
Tyr Pro Phe Asp Phe Gln Gly Ala Arg Ile
5 10

(2) INFORMATION FOR SEQ ID NO: 7

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 AMINO ACIDS

(B) TYPE: AMINO ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7

Lys Ser Asp Thr Gln Glu Thr Tyr Gly Ala Leu Asp Leu Gly Gly
Lys Ser Asp Thr Gln Glu Thr Tyr Gly Ala Leu Asp Leu Gly Gly
5 10 15
Ala Ser Thr Gln Val
20

(2) INFORMATION FOR SEQ ID NO: 8

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8

Lys Ser Asp Thr Gln Glu Thr Tyr Gly Ala Leu Asp Leu Gly Gly
5 10 15

Ala

bioRxiv preprint doi: <https://doi.org/10.1101/2022.05.10.485122>; this version posted May 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.